

Multivariate extensions of the Golden-Thompson inequality

Frank Hansen

June 22, 2014

Abstract

We study concave trace functions of several operator variables and formulate and prove multivariate generalisations of the Golden-Thompson inequality. The obtained results imply that certain functionals in quantum statistical mechanics have bounds of the same form as they appear in classical physics.

MSC2010 classification: 47A63.

Key words and phrases: Golden Thompson's inequality; multivariate trace inequality; concave trace function.

1 Introduction

The Golden-Thompson inequality, which is of importance in statistical mechanics and in the theory of random matrices, states that

$$\mathrm{Tr} e^{L+B} \leq \mathrm{Tr} e^L e^B$$

for arbitrary self-adjoint matrices L and B . It is known that there is no direct extension of this inequality to more operator variables, and there is an extensive literature investigating these matters, cf. [6, 1, 3] and the references therein.

We prove, among other statements, the following extension of the Golden-Thompson inequality. Consider $n \times m$ matrices H_1, \dots, H_k with

$$H_1^* H_1 + \dots + H_k^* H_k = 1_n.$$

Then the inequality

$$(1) \quad \text{Tr} \exp\left(L + \sum_{i=1}^k H_i^* B_i H_i\right) \leq \text{Tr} \exp(L) \sum_{i=1}^k H_i^* (\exp B_i) H_i$$

is valid for arbitrary self-adjoint $n \times n$ matrices L and B_1, \dots, B_k . This is, for $n = m$, the same bound as obtained when all the matrices commute. We are thus allowed to estimate partition functions or the Helmholtz function in quantum statistical mechanics and obtain bounds on the same form as they appear in classical physics.

We obtain, as a simple special case, the inequality

$$(2) \quad \text{Tr} \exp\left(L + \frac{1}{2}A + \frac{1}{2}B\right) \leq \text{Tr}(\exp L) \left(\frac{1}{2} \exp A + \frac{1}{2} \exp B\right)$$

valid for arbitrary self-adjoint matrices L , A and B . Notice that (2) reduces to the Golden Thompson inequality for $A = B$ and to convexity under the trace of the exponential function for $L = 0$. The inequality may thus be considered as an interpolation inequality between Golden-Thompson's inequality and Jensen's inequality. However, we cannot derive (2) from these special cases. If we first apply Golden-Thompson's inequality then we obtain

$$\text{Tr} \exp\left(L + \frac{1}{2}A + \frac{1}{2}B\right) \leq \text{Tr}(\exp L) \exp\left(\frac{1}{2}A + \frac{1}{2}B\right),$$

but this is insufficient to obtain (2) since L is arbitrary and the exponential function is not operator convex.

2 Preliminaries

The following lemma is both well-known and very useful. We include the proof for the benefit of the reader.

Lemma 2.1. *Let $\varphi: \mathcal{D} \rightarrow \mathcal{A}_{sa}$ be a map defined in a convex cone $\mathcal{D} \subseteq X$ of a Banach space X with values in the self-adjoint part of a C^* -algebra \mathcal{A} . If φ is Fréchet differentiable, convex and positively homogeneous then*

$$d\varphi(x)h \leq \varphi(h) \quad x, h \in \mathcal{D},$$

where $d\varphi(x)$ denotes the Fréchet differential of $\varphi(x)$.

Proof. Since

$$\begin{aligned}\varphi(x+th) &= (1+t)\varphi\left(\frac{1}{1+t}x + \frac{t}{1+t}h\right) \\ &\leq (1+t)\left(\frac{1}{1+t}\varphi(x) + \frac{t}{1+t}\varphi(h)\right) \\ &= \varphi(x) + t\varphi(h)\end{aligned}$$

for $0 \leq t \leq 1$, we obtain

$$\frac{\varphi(x+th) - \varphi(x)}{t} \leq \varphi(h) \quad 0 < t \leq 1$$

and thus $d\varphi(x)h \leq \varphi(h)$. **QED**

We refer to the monograph [2] for a general account of Fréchet differentiable mappings between Banach spaces.

The logarithm is operator monotone with Lebesgue measure as representing measure, thus

$$\log x = \int_0^\infty \left(\frac{1}{t+1} - \frac{1}{x+t} \right) dt \quad x > 0.$$

Since

$$x^{1/2}(x+h)^{-1}x^{1/2} = (1 + x^{-1/2}hx^{-1/2})^{-1} = 1 - x^{-1/2}hx^{-1/2} + o(h),$$

we derive that $dx^{-1}h = -x^{-1}hx^{-1}$. Notice that h may be arbitrary as we are not using the functional calculus. Consequently

$$d\log(x)h = \int_0^\infty (x+t)^{-1}h(x+t)^{-1} dt.$$

We have thus obtained the following integral expression

$$(3) \quad Q(x, h) = \text{Tr } h^* d\log(x)h = \text{Tr } \int_0^\infty h^*(x+t)^{-1}h(x+t)^{-1} dt.$$

It follows from the integral expression that $Q(x, h)$ is positively homogeneous in (x, h) . Lieb proved that it is a convex function in two variables [7, Theorem 3]. But this is a reflection of a quite general result. Zhang and the author

recently proved [5] that for a strictly increasing continuously differentiable function $f: (0, \infty) \rightarrow \mathbf{R}$ the form

$$(x, h) \rightarrow \text{Tr } h^* df(x)h \quad x > 0$$

is convex if and only if the derivative of f is operator convex and numerically decreasing.

We shall now use Lemma 2.1. Notice that $Q(x, h)$ is defined in the cone $\mathcal{D} = B(H)_+ \times B(H)$, where H is a finite dimensional Hilbert space. Thus

$$(4) \quad dQ(x, h)(y, k) \leq Q(y, k)$$

for positive definite x, y and arbitrary h, k .

We end this section by giving a new result for the form Q that will prove crucial in the rest of the paper.

Proposition 2.2. *Let X be an invertible contraction. Then*

$$Q(XAX^*, B) \leq Q(A, X^{-1}B(X^*)^{-1})$$

for positive definite A and arbitrary B .

Proof. We use the integral representation of the form Q and obtain

$$Q(XAX^*, B) = \text{Tr} \int_0^\infty B^*(XAX^* + t)^{-1} B(XAX^* + t)^{-1} dt.$$

Since X is a contraction we derive the inequality

$$\frac{1}{XAX^* + t} \leq \frac{1}{X(A + t)X^*} = (X^*)^{-1}(A + t)^{-1}X^{-1}.$$

Under the trace this inequality implies

$$\begin{aligned} Q(XAX^*, B) &\leq \text{Tr} \int_0^\infty B^*(X^*)^{-1}(A + t)^{-1}X^{-1}B(X^*)^{-1}(A + t)^{-1}X^{-1} dt \\ &= Q(A, X^{-1}B(X^*)^{-1}) \end{aligned}$$

which is the desired result. **QED**

3 Concave trace functions

Theorem 3.1. *Let H be a contraction. Then the trace function*

$$\varphi(A) = \text{Tr} \exp(H^*(\log A)H)$$

is concave in positive definite matrices.

Proof. We may without loss of generality assume that H is invertible. We calculate the first Fréchet differential

$$\begin{aligned} d\varphi(A)B &= \text{Tr} \, d\exp(H^*(\log A)H)(H^*(d\log(A)B)H) \\ &= \text{Tr} \exp(H^*(\log A)H)(H^*(d\log(A)B)H), \end{aligned}$$

where we used the identity $\text{Tr} \, df(A)B = \text{Tr} \, f'(A)B$ valid for differentiable functions. We then consider the following functions of the single operator variable A .

$$\begin{aligned} C &= H^*(\log A)H \\ D &= H^*(d\log(A)B)H = d_A(H^*(\log A)H)B = d_A(C)B \\ E &= H \exp(C)H^* = H \exp(H^*(\log A)H)H^* \\ G &= H \, d\exp(C)(D)H^* = d_C(H \exp(C)H^*)D = d_C(E)D. \end{aligned}$$

For clarity, we use the notation d_A to indicate Fréchet differentiation with respect to A of compound expressions. We proceed to calculate the second Fréchet differential

$$\begin{aligned} d^2\varphi(A)(B, B) &= d_A(d\varphi(A)B)B = d_A(\text{Tr} \exp(C)D)B \\ &= \text{Tr} \, d\exp(C)(D)D + \text{Tr} \exp(C)H^* d^2\log(A)(B, B)H \\ &= \text{Tr} \, d\exp(C)(D)D + \text{Tr} \, E \, d^2\log(A)(B, B). \end{aligned}$$

We recall [2] that φ is concave if and only if $d^2\varphi(A)(B, B) \leq 0$ for positive definite A and self-adjoint B . To evaluate the second term we apply the chain rule to the form $Q(A, B)$ and obtain

$$\begin{aligned} dQ(A, B)(a, b) &= d_1Q(A, B)a + d_2Q(A, B)b \\ &= \text{Tr} \, B \, d^2\log(A)(B, a) + \text{Tr} \, b \, d\log(A)B + \text{Tr} \, B \, d\log(A)b \end{aligned}$$

for positive definite A, a and self-adjoint B, b . The integral representation in (3) implies $\text{Tr } B \, d\log(A)b = \text{Tr } b \, d\log(A)B$, and we therefore obtain

$$dQ(A, B)(a, b) = \text{Tr } B \, d^2\log(A)(B, a) + 2\text{Tr } b \, d\log(A)B.$$

By using (4) we now obtain the inequality

$$\begin{aligned} \text{Tr } a \, d^2\log(A)(B, B) &= \text{Tr } B \, d^2\log(A)(B, a) \\ &= dQ(A, B)(a, b) - 2\text{Tr } b \, d\log(A)B \\ &\leq Q(a, b) - 2\text{Tr } b \, d\log(A)B \end{aligned}$$

for positive definite A, a and self-adjoint B, b . Since E is positive definite we may put $a = E$ and thus obtain

$$\begin{aligned} d^2\varphi(A)(B, B) &= \text{Tr } d\exp(C)(D)D + \text{Tr } E \, d^2\log(A)(B, B) \\ &\leq \text{Tr } d\exp(C)(D)D + Q(E, b) - 2\text{Tr } b \, d\log(A)B. \end{aligned}$$

By setting $b = G$ we then obtain

$$d^2\varphi(A)(B, B) \leq \text{Tr } d\exp(C)(D)D + Q(E, G) - 2\text{Tr } G \, d\log(A)B$$

for positive definite A and self-adjoint B . But since

$$\begin{aligned} \text{Tr } G \, d\log(A)B &= \text{Tr } H \, d\exp(C)(D)H^* \, d\log(A)B \\ &= \text{Tr } H^*(d\log(A)B)H \, d\exp(C)D \\ &= \text{Tr } d(H^*(\log A)H)B \, d\exp(C)D \\ &= \text{Tr } (d(C)B) \, d\exp(C)D \\ &= \text{Tr } D \, d\exp(C)D, \end{aligned}$$

we obtain

$$d^2\varphi(A)(B, B) \leq Q(E, G) - \text{Tr } D \, d\exp(C)D.$$

We now apply Proposition 2.2 and obtain

$$\begin{aligned} Q(E, G) &= Q(H \exp(C)H^*, G) \\ &\leq Q(\exp C, H^{-1}G(H^*)^{-1}) \\ &= Q(\exp C, d\exp(C)(D)) \\ &= \text{Tr } d\exp(C)(D) \, d\log(\exp C) \, d\exp(C)(D). \end{aligned}$$

However, since the inverse of the linear map $h \rightarrow d\exp(x)h$ is given by

$$d\exp(x)^{-1} = d\log(\exp x),$$

we realise that

$$d\log(\exp C) d\exp(C)(D) = d\exp(C)^{-1}(d\exp(C)D) = D.$$

Therefore,

$$Q(E, G) \leq \text{Tr } D d\exp(C)D$$

and thus $d^2\varphi(A)(B, B) \leq 0$ for positive definite A and self-adjoint B . This shows that φ is concave. **QED**

Corollary 3.2. *Consider $n \times m$ matrices H_1, \dots, H_k with*

$$H_1^*H_1 + \dots + H_k^*H_k \leq 1_n$$

where 1_n denotes the $n \times n$ unit matrix. Then the trace function

$$\varphi(A_1, \dots, A_k) = \text{Tr } \exp(H_1^*(\log A_1)H_1 + \dots + H_k^*(\log A_k)H_k)$$

is concave in k -tuples of positive definite $n \times n$ matrices.

Proof. We set

$$A = \begin{pmatrix} A_1 & 0 & \dots & 0 \\ 0 & A_2 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & A_k \end{pmatrix} \quad \text{and} \quad H = \begin{pmatrix} H_1 & 0 & \dots & 0 \\ H_2 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ H_k & 0 & \dots & 0 \end{pmatrix}$$

with zero matrices of suitable orders inserted and notice that H is a contraction. Furthermore,

$$H^*(\log A)H = \begin{pmatrix} \sum_{i=1}^k H_i^*(\log A_i)H_i & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}.$$

Thus

$$\text{Tr } \exp(H^*(\log A)H) = \text{Tr } \exp\left(\sum_{i=1}^k H_i^*(\log A_i)H_i\right) + (k-1)n$$

and the statement follows from Theorem 3.1. **QED**

Corollary 3.3. Consider $n \times m$ matrices H_1, \dots, H_k with

$$H_1^* H_1 + \dots + H_k^* H_k \leq 1_n$$

and a self-adjoint $n \times n$ matrix L . Then the trace function

$$\varphi(A_1, \dots, A_k) = \text{Tr} \exp\left(L + \sum_{i=1}^k H_i^* (\log A_i) H_i\right)$$

is concave in k -tuples of positive definite $n \times n$ matrices.

Proof. By appealing to continuity we may without loss of generality assume

$$H_1^* H_1 + \dots + H_k^* H_k < 1_n$$

and set $H_{k+1} = (1_n - (H_1^* H_1 + \dots + H_k^* H_k))^{1/2}$. Then H_{k+1} is positive definite and since

$$H_1^* H_1 + \dots + H_k^* H_k + H_{k+1}^2 = 1_n$$

we deduce from Corollary 3.2 that the trace function

$$\begin{aligned} & \varphi(A_1, \dots, A_k, A_{k+1}) \\ &= \text{Tr} \exp\left(H_1^* (\log A_1) H_1 + \dots + H_k^* (\log A_k) H_k + H_{k+1} (\log A_{k+1}) H_{k+1}\right) \end{aligned}$$

is concave in positive definite matrices. We keep A_{k+1} constant by setting

$$A_{k+1} = \exp(H_{k+1}^{-1} L H_{k+1}^{-1})$$

and the statement now follows. **QED**

Remark 3.4. Corollary 3.3 contains two celebrated theorems of Lieb. If we set $k = 1$ and $H = 1$ then the trace function

$$\varphi(A) = \text{Tr} \exp(L + \log A)$$

is concave in positive definite matrices, cf. [7, Theorem 6]. If H_1, \dots, H_k are chosen as square roots of positive numbers times the identity matrix then we obtain that the trace function

$$\varphi(A_1, \dots, A_k) = \text{Tr} \exp(L + p_1 \log A_1 + \dots + p_k \log A_k),$$

defined in positive definite matrices, is concave, where p_1, \dots, p_k are non-negative numbers with $p_1 + \dots + p_k \leq 1$, cf. [7, Corollary 6.1 (1)].

Corollary 3.5. *Let L be a fixed self-adjoint matrix, and let A_1, \dots, A_k be random self-adjoint matrices. Then the inequality*

$$\mathbb{E} \operatorname{Tr} \exp \left(L + \sum_{i=1}^k H_i^* A_i H_i \right) \leq \operatorname{Tr} \exp \left(L + \sum_{i=1}^k H_i^* (\log \mathbb{E} e^{A_i}) H_i \right)$$

holds for fixed matrices H_1, \dots, H_k with $H_1^ H_1 + \dots + H_k^* H_k \leq 1$.*

The result follows directly from Corollary 3.3 by applying Jensen's inequality, cf. also [9, Corollary 3.3]. A simple consequence is that

$$\mathbb{E} \operatorname{Tr} \exp \left(L + \frac{A_1 + \dots + A_k}{k} \right) \leq \operatorname{Tr} \exp \left(L + \frac{\log \mathbb{E} e^{A_1} + \dots + \log \mathbb{E} e^{A_k}}{k} \right)$$

for a fixed self-adjoint matrix L and random self-adjoint matrices A_1, \dots, A_k .

4 Multivariate trace inequalities

Lemma 4.1. *Consider $n \times m$ matrices H_1, \dots, H_k with*

$$H_1^* H_1 + \dots + H_k^* H_k = 1_n$$

and a self-adjoint $n \times n$ matrix L . Then we have the inequality

$$\begin{aligned} & \operatorname{Tr} \exp \left(L + \sum_{j=1}^k H_j^* (\log B_j) H_j \right) \\ & \leq \operatorname{Tr} \exp \left(L + \sum_{j=1}^k H_j^* (\log A_j) H_j \right) \sum_{i=1}^k H_i^* (d \log(A_i) B_i) H_i \end{aligned}$$

for positive definite $n \times n$ matrices A_1, \dots, A_k and B_1, \dots, B_k .

Proof. Since the trace function

$$\varphi(A_1, \dots, A_k) = \operatorname{Tr} \exp \left(L + \sum_{j=1}^k H_j^* (\log A_j) H_j \right)$$

is concave and (positively) homogeneous, we may apply Lemma 2.1 and obtain the inequality

$$d\varphi(A_1, \dots, A_k)(B_1, \dots, B_k) \geq \varphi(B_1, \dots, B_k).$$

By applying the chain rule for Fréchet differentials we then derive

$$\begin{aligned}
\varphi(B_1, \dots, B_k) &\leq \sum_{i=1}^k d_i \varphi(A_1, \dots, A_k) B_i \\
&= \sum_{i=1}^k \text{Tr} \, d \exp \left(L + \sum_{j=1}^k H_j^* (\log A_j) H_j \right) H_i^* (d \log(A_i) B_i) H_i \\
&= \sum_{i=1}^k \text{Tr} \, \exp \left(L + \sum_{j=1}^k H_j^* (\log A_j) H_j \right) H_i^* (d \log(A_i) B_i) H_i
\end{aligned}$$

and the statement follows. **QED**

Theorem 4.2. Consider $n \times m$ matrices H_1, \dots, H_k with

$$H_1^* H_1 + \dots + H_k^* H_k = 1_n.$$

Then we have the inequality

$$\text{Tr} \exp \left(L + \sum_{i=1}^k H_i^* B_i H_i \right) \leq \text{Tr} \exp(L) \sum_{i=1}^k H_i^* (\exp B_i) H_i$$

for arbitrary self-adjoint $n \times n$ matrices L and B_1, \dots, B_k .

Proof. Choose in Lemma 4.1 for $i = 1, \dots, k$ the matrix A_i as the identity matrix. Then the Fréchet differential $d \log(A_i) B_i = B_i$ and $\log A_i = 0$. The result then follows by replacing B_i with $\exp B_i$ for $i = 1, \dots, k$. **QED**

The above inequality is a direct generalisation of the Golden-Thompson inequality. Indeed, if we put $k = 1$ and take H_1 as the identity matrix then the inequality in Theorem 4.2 takes the form

$$\text{Tr} e^{L+B} \leq \text{Tr} e^L e^B,$$

cf. [4, 8, 7]. We may obtain other corollaries of Lemma 4.1.

Theorem 4.3. Consider $n \times m$ matrices H_1, \dots, H_k with

$$H_1^* H_1 + \dots + H_k^* H_k = 1_n.$$

Then we have the inequality

$$\text{Tr} \exp \left(\sum_{i=1}^k H_i^* (\log B_i - \log A_i) H_i \right) \leq \sum_{i=1}^k \text{Tr} H_i^* (d \log(A_i) B_i) H_i$$

for positive definite $n \times n$ matrices A_1, \dots, A_k and B_1, \dots, B_k .

Proof. The result follows by setting

$$L = -(H_1^*(\log A_1)H_1 + \cdots + H_k^*(\log A_k)H_k)$$

in Lemma 4.1. **QED**

If we in Theorem 4.3 put $k = 1$ and take H_1 as the unit matrix we obtain

$$\mathrm{Tr} \exp(\log B - \log A) \leq \mathrm{Tr} \, d\log(A)B = \mathrm{Tr} A^{-1}B$$

which is the Golden-Thompson inequality.

Furthermore, if A_i and B_i commute for $i = 1, \dots, k$ then Theorem 4.3 reduces to the inequality

$$\mathrm{Tr} \exp\left(\sum_{i=1}^k H_i^*(\log B_i - \log A_i)H_i\right) \leq \sum_{i=1}^k \mathrm{Tr} H_i^* B_i A_i^{-1} H_i$$

which is an expression of operator convexity of the exponential function under the trace.

Theorem 4.4. Consider $n \times m$ matrices H_1, \dots, H_k with

$$H_1^* H_1 + \cdots + H_k^* H_k = 1_n.$$

Then we have the inequality

$$\begin{aligned} & \mathrm{Tr} \exp\left(\sum_{i=1}^k H_i^*(\log B_i + \log C_i - \log A_i)H_i\right) \\ & \leq \mathrm{Tr} \exp\left(\sum_{i=1}^k H_i^*(\log C_i)H_i\right) \sum_{i=1}^k H_i^*(d\log(A_i)B_i)H_i \end{aligned}$$

for positive definite $n \times n$ matrices $A_1, \dots, A_k, B_1, \dots, B_k$ and C_1, \dots, C_k .

Proof. The result follows by setting

$$L = H_1^*(\log C_1 - \log A_1)H_1 + \cdots + H_k^*(\log C_k - \log A_k)H_k$$

in Lemma 4.1. **QED**

If we in Theorem 4.4 put $k = 1$ and take H_1 as the unit matrix we obtain

$$\mathrm{Tr} \exp(\log B + \log C - \log A) \leq \mathrm{Tr} C \, d\log(A)B,$$

which is the extended Golden-Thompson inequality. The extended Golden-Thompson inequality reduces to the Golden-Thompson inequality if A and B commute.

References

- [1] T. Ando and F. Hiai. Log majorization and complementary Golden-Thompson type inequalities. *Linear Algebra Appl.*, 197 (198):113–131, 1994.
- [2] T.M. Flett. *Differential Analysis*. Cambridge University Press, Cambridge, 1980.
- [3] P.J. Forrester and C.J. Thompson. The Golden-Thompson inequality: Historical aspects and random matrix applications. *J. Math. Phys.*, 55(023503), 2014.
- [4] S. Golden. Lower bounds for the Helmholtz function. *Phys. Rev. B*, 137(1127–1128), 1965.
- [5] F. Hansen and Z. Zhang. Characterisation of matrix entropies. *arXiv: 1402.2118*, pages 1–11, 2014.
- [6] F. Hiai and D. Petz. The Golden-Thomson trace inequality is complemented. *Linear Algebra Appl.*, 181:153–185, 1993.
- [7] E. Lieb. Convex trace functions and the Wigner-Yanase-Dyson conjecture. *Advances in Math.*, 11:267–288, 1973.
- [8] C.J. Thompson. Inequality with applications in statistical mechanics. *J. Math. Phys.*, 6:1812–1813, 1965.
- [9] J.A. Tropp. User-friendly tail bounds for sums of random variables. *Found Comput Math*, 12:389–434, 2012.

Frank Hansen: Institute for Excellence in Higher Education, Tohoku University, Japan.
Email: frank.hansen@m.tohoku.ac.jp.